DEPARTMENT OF ECONOMICS AND FINANCE

COLLEGE OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

Evaluating Macroeconomic Forecasts: A Concise Review of Some Recent Developments

Philip Hans Franses

Michael McAleer

Rianne Legerstee

WORKING PAPER

No. 12/2012

Department of Economics and Finance
College of Business and Economics
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

WORKING PAPER No. 12/2012

Evaluating Macroeconomic Forecasts: A Concise Review of Some Recent Developments

Philip Hans Franses[1], Michael McAleer[2], Rianne Legerstee[3]

8 June 2012

Abstract: Macroeconomic forecasts are frequently produced, widely published, intensively discussed and comprehensively used. The formal evaluation of such forecasts has a long research history. Recently, a new angle to the evaluation of forecasts has been addressed, and in this review we analyse some recent developments from that perspective. The literature on forecast evaluation predominantly assumes that macroeconomic forecasts are generated from econometric models. In practice, however, most macroeconomic forecasts, such as those from the IMF, World Bank, OECD, Federal Reserve Board, Federal Open Market Committee (FOMC) and the ECB, are typically based on econometric model forecasts jointly with human intuition. This seemingly inevitable combination renders most of these forecasts biased and, as such, their evaluation becomes non-standard. In this review, we consider the evaluation of two forecasts in which: (i) the two forecasts are generated from two distinct econometric models; (ii) one forecast is generated from an econometric model and the other is obtained as a combination of a model and intuition; and (iii) the two forecasts are generated from two distinct (but unknown) combinations of different models and intuition. It is shown that alternative tools are needed to compare and evaluate the forecasts in each of these three situations. These alternative techniques are illustrated by comparing the forecasts from the (econometric) Staff of the Federal Reserve Board and the FOMC on inflation, unemployment and real GDP growth. It is shown that the FOMC does not forecast significantly better than the Staff, and that the intuition of the FOMC does not add significantly in forecasting the actual values of the economic fundamentals. This would seem to belie the purported expertise of the FOMC.

Keywords: Macroeconomic forecasts, econometric models, human intuition, biased forecasts, forecast performance, forecast evaluation, forecast comparison.

JEL Classifications: C22, C51, C52, C53, E27, E37.

Acknowledgements: The authors wish to thank Les Oxley, Chia-Lin Chang, an Associate Editor and two anonymous referees for detailed and helpful comments and suggestions. The second author wishes to acknowledge the financial support of the Australian Research Council, National Science Council, Taiwan, and the Japan Society for the Promotion of Science.

[1] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam

[2] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam; Tinbergen Institute, The Netherlands; Department of Quantitative Economics, Complutense University of Madrid; Institute of Economic Research, Kyoto University

[3] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam; Tinbergen Institute, The Netherlands

* Corresponding Author:michael.mcaleer@gmail.com

WORKING PAPER No. 12/2012

Evaluating Macroeconomic Forecasts: A Concise Review of Some Recent Developments

1. Introduction.

Macroeconomic forecasts are frequently produced, widely published, intensively discussed and comprehensively used. The formal evaluation of such forecasts has a long research history. There are many studies on the design of appropriate evaluation criteria (see, for example, Chong and Hendry (1986), Granger and Newbold (1986), Elliott and Timmermann (2008), and various chapters in Clements and Hendry (2002)). There has also been considerable discussion about the proper use of data, as macroeconomic data are frequently revised over time. Thus, the important question arises as to which vintage of data is the most relevant. There is also a considerable literature on alternative combinations of forecasts. Indeed, it may well be that combined forecasts outperform individual forecasts (see Timmermann (2006) for a recent survey).

The situation to be reviewed in this paper addresses a different aspect of macroeconomic forecasts, which may be presented as follows. The analyst has two (or more) forecasts and it is unknown how these forecasts were constructed. In order to keep the notation simple, we assume that each forecast is constructed as follows. The forecaster has access to an econometric model-based forecast (*MF*) and combines this with personal intuition (*I*), as follows:

$$Forecast = \alpha MF + \beta I$$

For each forecaster, the values of *MF*, α, *I* and β are known. In what follows, we assume the structural stability of an econometric forecasting model. A second forecaster also has a way of arriving at a final forecast. The analyst does not know any of these values, having only the two forecasts[1].

[1] In business forecasting, similar situations can occur when a statistical forecast can be adjusted by a manager. Due to the increased availability of relevant data, it is only recently that several studies have begun to focus on the behavior of managers when they receive these statistical forecasts. Indeed, in business forecasting, it is also acknowledged that analysts do need to know what it is that managers do before they can evaluate the quality of their adjusted forecasts. Recent studies are Fildes et al. (2009), Franses and Legerstee (2009), and Eroglu and Croxton (2010).

This paper is concerned with comparing the two resultant forecasts. Note that, when $\alpha = 0$, the final forecast is fully based on intuition. When $\beta = 0$ and $\alpha = 1$, the final forecast is based fully on an econometric model. This second case is the typical case studied in most econometrics textbooks.

Forecasts of the above type are provided, for example, by the Survey of Professional Forecasters (SPF), delivering the mean of forecasts reported by various experts. Indeed, it is not likely that all their forecasts are based on econometric models. Recently, Franses et al. (2011) documented that all forecasts from the Netherlands Bureau for Economic Policy Analysis (CPB) are the weighted sum of an econometric model forecast (based on a model comprising 2500 equations) and intuition. In the same spirit, it is likely that forecasts reported by, among others, the IMF, World Bank and the OECD are almost certainly obtained in a similar way.

In this concise review, we address the issue of evaluating macroeconomic forecasts when they might be based only partly on econometric models, in a way that is unknown to the analyst. The main focus of some recent developments in this area is that the analyst somehow has to disentangle the replicable from the non-replicable components of these forecasts, whereby the analyst can use a publicly available information set. This replicable part then mimics an econometric model that might have been used. The remainder of the forecasts, namely the non-replicable component, is associated with intuition, as it cannot be replicated by the analyst. As forecasters can and do incorporate the forecasts provided by other forecasters before presenting their own, the publicly available information set would typically also contain previously published forecasts. When formally comparing the forecasts, it is necessary to use alternative econometric tools as the variables of interest will turn out to be generated regressors, which contain estimation error.

The outline of the remainder of the paper is as follows. In Section 2 we address three simple cases, which can naturally be extended in various directions. Consider two forecasts that might be generated from: (i) two distinct econometric models; (ii) an econometric model and a combination of model and intuition; and (iii) two distinct (but unknown) combinations of model and intuition. It is shown that, in each situation, alternative tools are needed to compare and evaluate the forecasts. In Section 3 we illustrate the alternative cases by comparing the forecasts from the Federal Reserve Board and the FOMC on inflation, unemployment and real GDP growth. It is shown that each of the three situations can lead to significantly different evaluation outcomes.

2. Model Specifications.

This section reviews three different cases concerning two macroeconomic forecasts, wherein the analyst has to evaluate their relative quality and performance under different assumptions regarding the econometric model and intuition.

.

2.1 Forecasts from two econometric models.

Consider the variable of interest, X_t, and the availability of two sets of one-step-ahead forecasts, $F_{1,t}$ and $F_{2,t}$, for the sample $t = n+1, n+2, ..., n+N$. When the forecasts are based on linear econometric models, these models may be given as

$$X_t = W_{1,t}\beta_1 + \varepsilon_{1,t} \quad (1)$$

$$X_t = W_{2,t}\beta_2 + \varepsilon_{2,t} \quad (2)$$

where the information sets for the two econometric models are given, respectively, as $W_{1,t}$ and $W_{2,t}$, each including a column of ones. When OLS is used to estimate the unknown parameters, the unbiased forecasts are given as

$$F_{1,t} = W_{1,t}\hat{\beta}_1 \quad (3)$$

$$F_{2,t} = W_{2,t}\hat{\beta}_2 \quad (4)$$

In practice, it is quite likely that only the outcomes, $F_{1,t}$ and $F_{2,t}$, are available to the analyst, but the information sets, $W_{1,t}$ and $W_{2,t}$, are not made available. Let us assume that the analyst can resort to the publicly available information set, W_t, which can include both $W_{1,t}$ and $W_{2,t}$ ..

When it is known that $W_{1,t}$ nests $W_{2,t}$, the techniques developed in Clark and McCracken (2001) are useful. If the models are non-nested, one can rely on, for example, the Diebold and Mariano (1995) test (see also West (1996)).

An alternative simple method that might be used when little is known about $W_{1,t}$ and $W_{2,t}$ relies on the auxiliary regression:

$$X_t = \alpha_1 F_{1,t} + \alpha_2 F_{2,t} + \xi_t \quad (5)$$

This regression is also at the heart of combinations of forecasts (see Timmermann, 2006). Regression (5) can be used to examine whether each of the forecasts adds significantly to the other forecast. If so, then one may want to combine the two forecasts, with the parameters in (5) being used as weights.

2.2 One forecast from a model, the other from a combination of model and intuition.

A second case is the following. Suppose that the second forecast $F_{2,t}$ is partly based on a model, but also partly based on the first forecast, $F_{1,t}$, and on intuition, that is:

$$F_{1,t} = W_{1,t}\hat{\beta}_1 \tag{6}$$

$$F_{2,t} = W_{2,t}\hat{\beta}_2 + \gamma F_{1,t} + \eta_{2,t} \tag{7}$$

where $\eta_{2,t}$ denotes the intuition included in the second forecast. When $W_{2,t}\hat{\beta}_2$ in (7) is the outcome of some econometric model, then that part of (7) is unbiased, but the two added terms, $\gamma F_{1,t} + \eta_{1,t}$, may cause bias. Evidence for the presence of bias in macroeconomic forecasts is presented in Batchelor (2007), among others.

It is evident that now the regression

$$X_t = \alpha_1 F_{1,t} + \alpha_2 F_{2,t} + \xi_t \tag{8}$$

cannot be used in a straightforward manner as the forecast $F_{2,t}$ contains $F_{1,t}$. Franses et al. (2009) and Chang et al. (2011) propose the auxiliary regression

$$F_{2,t} = W_t\beta + \pi F_{1,t} + \varsigma_t \tag{9}$$

in order to estimate

$$X_t = \alpha_1 F_{1,t} + \alpha_2 \hat{F}_{2,t} + \xi_t \tag{10}$$

where $\hat{F}_{2,t} = W_t\hat{\beta} + \hat{\pi} F_{1,t}$ is obtained from (9).

As $\hat{F}_{2,t}$ is a generated regressor, the econometric analysis of (10) is non-standard. When the difference between $\hat{F}_{2,t}$ and $F_{2,t}$ is viewed as a measurement error, the covariance matrix of ξ_t in (10) is not proportional to the identity matrix, so that ξ_t is serially correlated and heteroskedastic. However, as Franses et al. (2009) demonstrate, OLS estimation of the parameters in (10) can nevertheless be consistent and efficient.

Franses et al. (2009) establish the conditions under which OLS estimation of the parameters in a more general version of (10) is efficient by appealing to Kruskal's Theorem, which is necessary and sufficient for OLS to be efficient (see McAleer and McKenzie (1991), Fiebig et al. (1992) and McAleer (1992) for further details). In the context of OLS estimation of (10), the necessary and sufficient conditions for OLS to be efficient will be satisfied either if the variables used to obtain the forecast $F_{1,t}$ are contained in the information set of the forecast $F_{2,t}$, or are orthogonal to the variables in the information set of $F_{2,t}$. Of the two alternative necessary and sufficient conditions, it is more likely that the former condition will hold.

It was also shown in Franses et al. (2009) that, if the incorrect downward biased OLS standard errors are used, then the incorrect OLS t-ratios will be biased upward. Therefore, they suggest that the correct OLS covariance matrix in (10) should be estimated consistently using the Newey-West HAC standard errors (see also Smith and McAleer (1994)). Finally, Franses et al. (2009) and Chang et al. (2011) propose using GMM as an alternative to OLS.

2.3 Both forecasts as distinct combinations of model and intuition.

A third case, which may be the most likely to occur in practice, is where both forecasts are distinct combinations of model and intuition. To the analyst, the nature of this combination is unknown. It is also most likely that there is no documentation regarding any such intuition.

Franses et al. (2011) document that, at the Netherlands Bureau for Economic Policy Analysis (CPB), detailed records are retained of the size of any changes, but not of the motivation for the size of any changes. Hence, one may presume that the analyst has forecasts that might be generated as follows:

$$F_{1,t} = W_{1,t}\hat{\beta}_1 + \eta_{1,t} \tag{11}$$

$$F_{2,t} = W_{2,t}\hat{\beta}_2 + \gamma F_{1,t} + \eta_{2,t} \tag{12}$$

where $\eta_{1,t}$ and $\eta_{2,t}$ are intuition, and where we again assume that forecaster 2 has knowledge of the other forecast. If this is not the case, one can impose the restriction $\gamma = 0$ in (12).

In order to evaluate the relative merits of these two forecasts, one would then run the regressions:

$$F_{1,t} = W_t\beta_1 + \varsigma_{1,t} \tag{13}$$

$$F_{2,t} = W_t\beta_2 + \varsigma_{2,t} \tag{14}$$

First, as in case 2, one may consider the auxiliary regression:

$$X_t = \alpha_1 \hat{F}_{1,t} + \alpha_2 \hat{F}_{2,t} + \xi_t \qquad (15)$$

where $\hat{F}_{2,t}$ and $\hat{F}_{1,t}$ are obtained from (13) and (14), respectively. Again, OLS is consistent, but HAC standard errors are required for valid inferences to be drawn.

Second, one may also examine what the forecasts might add to what an analyst can do using publicly available information, and this would be based on the auxiliary regression:

$$X_t = W_t\beta + \alpha_1 \hat{\varsigma}_{1,t} + \alpha_2 \hat{\varsigma}_{2,t} + \xi_t \qquad (16)$$

where $\hat{\varsigma}_{1,t}$ and $\hat{\varsigma}_{2,t}$ are the estimated residuals from (13) and (14), respectively. As W_t denotes publicly available information, the regression in (16) informs whether the intuition (which is not observable, but rather is estimated) of forecaster 1 and/or of forecaster 2 adds any significant value to the final forecast. For example, if the estimate of α_1 in (16) is significant, then one can conclude that the intuition of forecaster 1 adds to forecast accuracy of the forecast based solely on W_t.

3. Evaluating FOMC and Staff Forecasts.

In this section we evaluate empirically the above three cases using data that were recently analyzed in Romer and Romer (2008). In their study, they compare Staff and FOMC forecasts, and their starting point is case 1 in Section 2. In this section, we examine if a change in assumptions regarding how the forecasts were obtained, namely cases 2 and 3, can materially change the conclusions reached in Romer and Romer (2008) regarding the superiority of Staff versus FOMC forecasts.

The variables of interest, X_t, in Romer and Romer (2008) are the (real time) inflation rate, unemployment rate, and the real growth rate. The data for the empirical analysis are described in Romer and Romer (2008, pp. 230-231), and are available in an appendix on the AEA website (http://www.aeaweb.org/articles/issues_datasets.php). As discussed in Romer and Romer (2008, pp. 230-231), the FOMC prepares forecasts in February and July each year. The February forecasts for inflation and the growth rate are for the four quarters ending in the fourth quarter of the current year, and the unemployment rate forecast is for the fourth quarter of the current year. The July forecasts are for the same variables for both the current and next year. The sample is from 1979 to 2001, with 22 February forecasts and 46 July forecasts, giving a total of 68 observations.

[Insert Figures 1-3 about here]

The actual inflation rate, unemployment rate and real growth rate, as well as the corresponding staff and FOMC forecasts, are shown in Figures 1-3, respectively. It is clear that the staff and FOMC forecasts are very similar, but it is also clear that they are not particularly close to the actual rates they are forecasting, which raises the question as to how much better these forecasts are relative to those that an analyst could make based on publicly available information. The similarity in the two sets of forecasts is supported by the correlations in Table 1 between the staff and FOMC forecasts, which are obviously very close to each other.

[Insert Table 1 about here]

The similarity in forecast performance is also shown in Table 2, which reports the mean and median squared prediction errors for the staff and FOMC forecasts for the three variables. The staff seems slightly better than the FOMC in forecasting the inflation rate, the reverse holds in forecasting the real growth rate, and it is too close to call for the unemployment rate, with the staff only slightly better (worse) than the FOMC in terms of the mean (median) squared prediction error. In terms of forecasting performance, therefore, it would be fair to call the outcome a tie. The Diebold-Mariano test supports this conclusion as the test does not indicate significant differences.

[Insert Table 2 about here]

Case 1: *Assume that Staff and FOMC forecasts are based purely on econometric models*.
Romer and Romer (2008) assume that Case 1 prevails in this situation, and they run the regression:

$$X_t = \mu + \alpha_1 S_t + \alpha_2 P_t + \xi_t \tag{17}$$

where S denotes the Staff forecast and P denotes the Policymaker (that is, FOMC) forecasts. In terms of formal tests of the forecasting performance of the staff and the FOMC, the OLS and GMM estimates of equation (17) are given in Table 3. When Case 2 would be the real situation and Case 1 is assumed, then OLS is inconsistent and the forecast is not MSE optimal, while GMM is consistent. For the instrument list for GMM, we use the one-period lagged values of inflation, unemployment rate and real growth rate (except for the case of real growth, where only the second lag is used for a better fit). We experimented with alternative sets of instruments, but the qualitative results did not change significantly. We also examined recursive parameter estimates and recursive residuals, but we obtained no indication of structural breaks.

[Insert Table 3 about here]

For each variable, the first line reports the OLS results (which could be inconsistent in case 2), and the second line gives the GMM results. The OLS estimates correspond to those in Table 1 in Romer and Romer (2008), where it was inferred that the staff forecasts dominated those of the FOMC for inflation and the unemployment rate, though not for the real growth rate. It is instructive that the GMM estimates indicate that the staff is better than the FOMC in forecasting inflation, but not in forecasting the unemployment rate or the growth rate, where the effects of both the staff and FOMC forecasts are insignificant.

Although the OLS and GMM estimates of the coefficients are markedly different, it is worth noting that the sums (with estimated standard errors in parentheses) of the estimated staff and FOMC marginal effects are very similar, namely 1.00 (0.39) and 1.13 (1.44) for inflation, 0.94 (0.38) and 1.01 (1.79) for the unemployment rate, and 0.88 (0.49) and 1.19 (3.69) for the growth rate, which suggests that the estimates are economically meaningful. In this sense, the sum of the parts would seem to be greater than the whole.

Case 2: *Let the FOMC forecast be created after the Staff forecast is published, and assume that the Staff forecast is based on an econometric model*.
In this case we assume that (6) and (7) are useful, and are expressed as

$$S_t = W_{1,t}\hat{\beta}_1 \tag{18}$$

$$P_t = W_{2,t}\hat{\beta}_2 + \gamma S_t + \eta_t^p \tag{19}$$

This says that the Staff use an unknown econometric model, while the FOMC has a model, but also relies on the Staff forecasts and unobserved intuition, η_t^p. As analysts, we do not observe the information sets $W_{1,t}$ and $W_{2,t}$, and we do not know η_t^p. Hence, we rely on an auxiliary regression, as in (9), in order to calculate $\hat{P}_t$, which can be expressed as:

$$P_t = W_t\beta + \gamma S_t + \varsigma_t \qquad (20)$$

where, for W_t, we include one-period lagged values of inflation, unemployment rate and real growth rate to be consistent with the situation in case 1.

[Insert Table 4 about here]

The OLS estimates of equation (20), where (A) concerns the full model and (B) the case where $\beta = 0$, are given in Table 4. For purposes of estimating (20) (A), OLS is efficient and the forecast is MSE optimal, but OLS is inconsistent and the forecast is not MSE optimal for estimating (20) (B).

In the absence of additional variables other than the Staff forecasts, the inconsistent OLS estimates for (20) (B) might seem to suggest that the effect of the Staff forecast on the FOMC forecast is very close to unity for all three variables. However, the inclusion of additional variable available to the forecasters of the FOMC expertise, as approximated by one-period lagged inflation, unemployment and real growth rates, shows that the effect of the Staff forecast, while remaining significant, is considerably less. The F test of the joint significance of what FOMC adds to the Staff forecasts makes it clear it does matter, and significantly so, in obtaining the forecast P. In short, the FOMC uses information that is statistically significant.

[Insert Table 5 about here]

The empirical performance of the Staff and FOMC forecasts (after de-biasing) are compared in Table 5. The auxiliary regression is

$$X_t = \mu + \alpha_1 S_t + \alpha_2 \hat{P}_t + \xi_t \qquad (21)$$

where $\hat{P}_t$ is obtained from (20). Although OLS is efficient and the forecast is MSE optimal for equation (21), the standard errors are not proportional to the identity matrix, so the Newey-West HAC standard errors are calculated. The Staff is seen to dominate the FOMC for the inflation rate, but both the Staff and FOMC forecasts are insignificant for the unemployment and real growth rates. Although the goodness of fit of the OLS estimates in Tables 3 and 5 are virtually identical, the corresponding coefficient estimates are markedly different. However, the sums of the estimated staff and FOMC marginal effects in Table 5 are very similar to their OLS counterparts in Table 3, at 1.01 (0.55), 0.95 (0.71) and 0.98 (1.17) for inflation, unemployment rate and real growth rate, respectively, which suggest that the estimates are economically meaningful.

In summary, in a comparison with the Staff forecasts, the use of FOMC forecasts, as in Cases 1 or 2, yield considerably different empirical results. It can be seen clearly that the FOMC does not forecast well, but the same can be said about the Staff!

Case 3: *Assume that both forecasts are based on distinct combinations of model and intuition*.

In this situation, which seems most likely to hold in practice, we assume that (11) and (12) hold, which means that we run the auxiliary regressions:

$$S_t = W_t \beta_1 + \varsigma_{1,t} \tag{22}$$

$$P_t = W_t \beta + \gamma S_t + \varsigma_{2,t} \tag{23}$$

Next, we consider the auxiliary regression:

$$X_t = \mu + \alpha_1 \hat{S}_t + \alpha_2 \hat{P}_t + \xi_t \tag{24}$$

where $\hat{S}_t$ and $\hat{P}_t$ are obtained as the fitted values from (22) and (23), respectively.

[Insert Table 6 about here]

In Table 6, we report the OLS estimates of the parameters in (24), with the HAC standard errors. The evidence from this table demonstrates clearly that the Staff forecasts outperform the FOMC forecasts for all three variables, as the staff forecasts are significant whereas the FOMC forecasts are not.

The final situation that is of interest is to see whether the Staff and FOMC forecasts contain any unobservable intuition that might significantly add to what an analyst could achieve using publicly available information. We consider

$$X_t = W_t \beta + \alpha_1 \hat{\varsigma}_{1,t} + \alpha_2 \hat{\varsigma}_{2,t} + \xi_t \tag{25}$$

and report the estimates in Table 7, where it is found that the Staff intuition is significant for all three variables, whereas the FOMC intuition is not.

[Insert Table 7 about here]

These results are consistent with the findings in Table 6, namely that the intuition contained in the FOMC forecasts does not add significantly, whereas the intuition contained in the Staff forecasts does add significantly, in forecasting actual values of all three economic fundamentals. These results regarding intuition would seem to belie the purported expertise of the FOMC.

4. Conclusion.

The purpose of the paper was to provide a concise review of the evaluation of macroeconomic forecasts using alternative combinations of econometric model and intuition, which is the non-replicable component of forecasts.

In the empirical illustration, which was concerned with a comparison of the forecasts provided by the Federal Reserve Board's Staff and the Federal Open Market Committee (FOMC), it could safely be concluded that the FOMC did not add significantly to the forecasts of inflation, unemployment rate and the real growth rate, in comparison with the Staff.

Moreover, when we regressed the inflation rate, unemployment rate and the real growth rate on one-period lags of these three variables, we obtained mean squared prediction errors of 0.64, 0.25, and 1.31 respectively, while the median squared prediction errors are 0.03, 0.07 and 0.23. Hence, the analyst with simple forecasting tools could outperform both the Staff and the FOMC.

Table 7 suggested that the analyst could benefit from the intuition contained in the Staff forecasts, but not from the intuition in the FOMC forecasts, which would seem to belie the purported expertise of the FOMC.

This review concerned the situation that seems to prevail in practice. It is rarely found that macroeconomic forecasts are based on model outcomes only. When evaluating these forecasts, one can then not rely entirely on standard tools, as the added intuition may render the final forecasts biased. We evaluated some recent developments in this relatively new area, but it can safely be said that there are further developments to come.

Further work in this area may refer to the recent literature on forecasting inflation and other macroeconomic variables using model that allow parameters to change over time (see, for example, Stock and Watson (2003)). Furthermore, in our empirical analysis we only used three instruments, which were conveniently available at the same sampling frequency as the data of interest. In other settings, one could have forecasts created at the monthly level, and then would not necessarily have to rely on a few specific instruments. Indeed, one could then use multivariate techniques, such as principal components, to summarize a wealth of variables (see, for example, Heij et al. (2011) and the citations given therein). These techniques would prevent us from having to make specific choices for the inclusion of instruments as all available variables can be included automatically. Finally, one can elaborate on the small sample properties of the methods proposed in this paper, and one may also allow for time-varying parameters in order to allow for potential structural breaks in the data or in the model parameters.

Table 1

Correlations between Staff Forecasts and FOMC Forecasts

Variable	Correlation
Inflation	0.99
Unemployment	0.99
Real growth	0.97

Note: The sample is from 1979 to 2001, with 22 February forecasts and 46 July forecasts, giving a total of 68 observations.

Table 2

Comparison of Staff Forecasts and FOMC Forecasts

	Squared Prediction Errors			
	Mean		Median	
Variable	Staff	FOMC	Staff	FOMC
Inflation	0.71	0.89	0.19	0.28
Unemployment	0.54	0.57	0.16	0.15
Real growth	2.10	1.99	1.22	1.04

Note: The sample is from 1979 to 2001, with 22 February forecasts and 46 July forecasts, giving a total of 68 observations.

Table 3 (Case 1)

Comparison of Staff and FOMC Forecasts in Predicting Actual Values (standard errors are in parentheses)

Estimation method	Intercept	Staff (S_t)	FOMC (P_t)	R^2
		Inflation		
OLS	-0.20	1.10**	-0.10	0.86
	(0.22)	(0.39)	(0.37)	
GMM	-0.26	4.77**	-3.64	0.64
	(0.34)	(2.32)	(2.26)	
		Unemployment		
OLS	0.26	0.97*	-0.03	0.79
	(0.41)	(0.38)	(0.40)	
GMM	-0.37	3.41	-2.40	0.64
	(0.76)	(2.78)	(2.87)	
		Real growth		
OLS	0.43	0.25	0.63	0.44
	(0.36)	(0.49)	(0.52)	
GMM	-0.22	1.70	-0.51	0.31
	(0.83)	(3.61)	(3.42)	

Notes: The regression model is

$$X_t = \mu + \alpha_1 S_t + \alpha_2 P_t + \xi_t,$$

which is equation (1) in Romer and Romer (2008)), and equation (17) in the paper. The OLS estimates correspond to those in Table 1 of Romer and Romer (2008). The instrument list uses the one-period lagged values of inflation, unemployment rate and real growth (except for the case of real growth, where only lag 2 is used). * and ** denote significance at the 5% and 1% levels, respectively.

Table 4 (Case 2)

Auxiliary regressions to de-bias the FOMC forecasts
(standard errors are in parentheses)

	Inflation		Unemployment		Real growth	
Variables	(A)	(B)	(A)	(B)	(A)	(B)
Intercept	-0.18 (0.16)	0.01 (0.07)	-0.00 (0.13)	0.19 (0.12)	-0.22 (0.20)	0.28 (0.08)
Staff Forecast, S_t	0.91^{**} (0.06)	1.03^{**} (0.02)	0.77^{**} (0.06)	0.96^{**} (0.02)	0.86^{**} (0.04)	0.93^{**} (0.03)
P_{t-1}	0.38^{**} (0.12)		0.32^{**} (0.12)		0.33^{**} (0.12)	
S_{t-1}	-0.26^{*} (0.13)		-0.14 (0.12)		-0.19 (0.11)	
Inflation_{t-1}	-0.03 (0.04)		-0.00 (0.02)		0.02 (0.03)	
$\text{Unemployment}_{t-1}$	0.04 (0.03)		0.04 (0.05)		0.03 (0.03)	
Real growth_{t-1}	0.01 (0.02)		0.01 (0.02)		0.02 (0.03)	
R^2	0.99	0.98	0.98	0.98	0.96	0.94
F test		4.86^{**}		5.79^{**}		5.87^{**}

Notes: The regression equation correlates P_t and S_t through

$$P_t = W_t\beta + \gamma S_t + \varsigma_t$$

which is equation (20) in the paper. * and ** denote significance at the 5% and 1% levels, respectively.

Table 5 (Case 2)

Comparison of Staff and FOMC Forecasts in Predicting Actual Values: Staff forecasts are based on an econometric model, and FOMC forecasts are based on Staff forecasts, other variables and intuition (standard errors are in parentheses)

Estimation method	Intercept	Staff (S_t)	FOMC (P_t)	R^2
		Inflation		
OLS	-0.20	1.89**	-0.88	0.85
(HAC)	(0.25)	(0.55)	(0.56)	
		Unemployment		
OLS	0.22	0.80	0.15	0.79
(HAC)	(0.67)	(0.71)	(0.71)	
		Real growth		
OLS	0.10	-0.28	1.26	0.45
(HAC)	(0.48)	(1.07)	(1.06)	

Notes: The regression model is

$$X_t = a + \delta_0 S_t + \beta \hat{P}_t + \xi_t ,$$

which is equation (21) in the paper. The Newey-West HAC standard errors are given in parentheses. ** denotes significance at the 5% level.

Table 6

Comparison of Staff and FOMC Forecasts in Predicting Actual Values: Staff forecasts are based on an econometric model and intuition, and FOMC forecasts are based on Staff forecasts, other variables and intuition (HAC standard errors are in parentheses)

Estimation method	Intercept	Staff (S_t)	FOMC (P_t)	R^2
		Inflation		
OLS	-0.34	0.58^{*}	0.43	0.85
(HAC)	(0.29)	(0.27)	(0.23)	
		Unemployment		
OLS	-0.13	0.80^{**}	0.20	0.82
(HAC)	(0.67)	(0.20)	(0.14)	
		Real growth		
OLS	-0.95	1.16^{**}	0.30	0.62
(HAC)	(0.56)	(0.18)	(0.21)	

Notes: The regression model is

$$X_t = \mu + \alpha_1 \hat{S}_t + \alpha_2 \hat{P}_t + \xi_t ,$$

which is equation (24) in the paper. The Newey-West HAC standard errors are given in parentheses. *,** denotes significance at the 1% and 5% levels, respectively..

Table 7

Comparison of Staff and FOMC Forecasts in Predicting Actual Values: Intuition is added to the lagged variables (chosen by the analyst) (parameter estimates for lagged inflation, lagged unemployment and lagged growth are not reported) (HAC standard errors are in parentheses)

Estimation method	Intuition of Staff (S_t)	Intuition of FOMC (P_t)	R^2
	Inflation		
OLS	0.58*	0.25	0.87
(HAC)	(0.24)	(0.48)	
	Unemployment		
OLS	0.32**	-0.19	0.90
(HAC)	(0.10)	(0.40)	
	Real growth		
OLS	0.29*	0.45	0.65
(HAC)	(0.15)	(0.62)	

Notes: The regression model is

$$X_t = W_t\beta + \alpha_1\hat{\varsigma}_{1,t} + \alpha_2\hat{\varsigma}_{2,t} + \xi_t,$$

which is equation (25) in the paper. The Newey-West HAC standard errors are given in parentheses. *,** denotes significance at the 1% and 5% levels, respectively

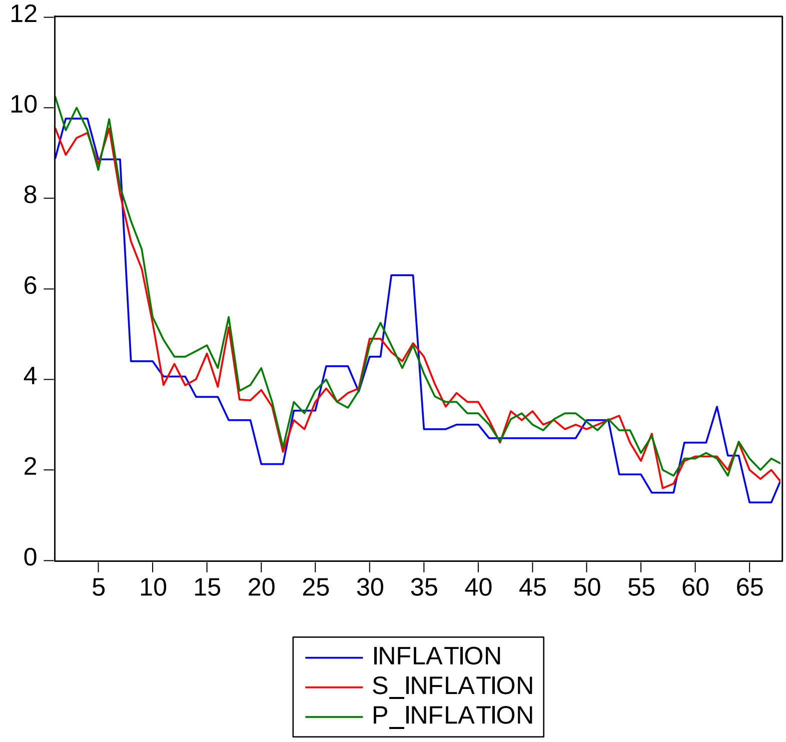

Figure 1

Inflation rate, Staff forecasts (S_inflation) and FOMC forecasts (P_inflation)

The sample is from 1979 to 2001, with 22 February forecasts and 46 July forecasts (1979 is observation 1 and 2001 is observation 68).

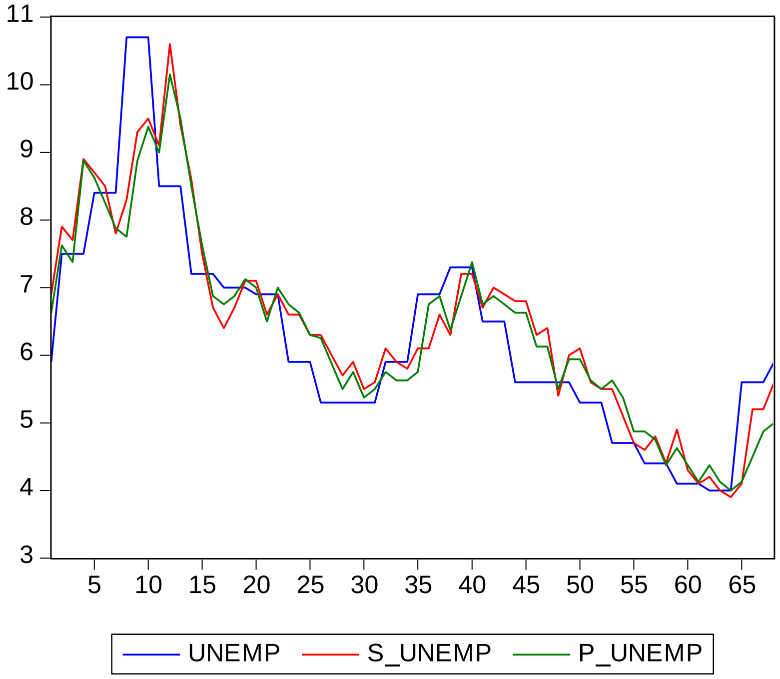

Figure 2

Unemployment rate, Staff forecasts (S_unemp) and FOMC forecasts (P_unemp)

The sample is from 1979 to 2001, with 22 February forecasts and 46 July forecasts (1979 is observation 1 and 2001 is observation 68).

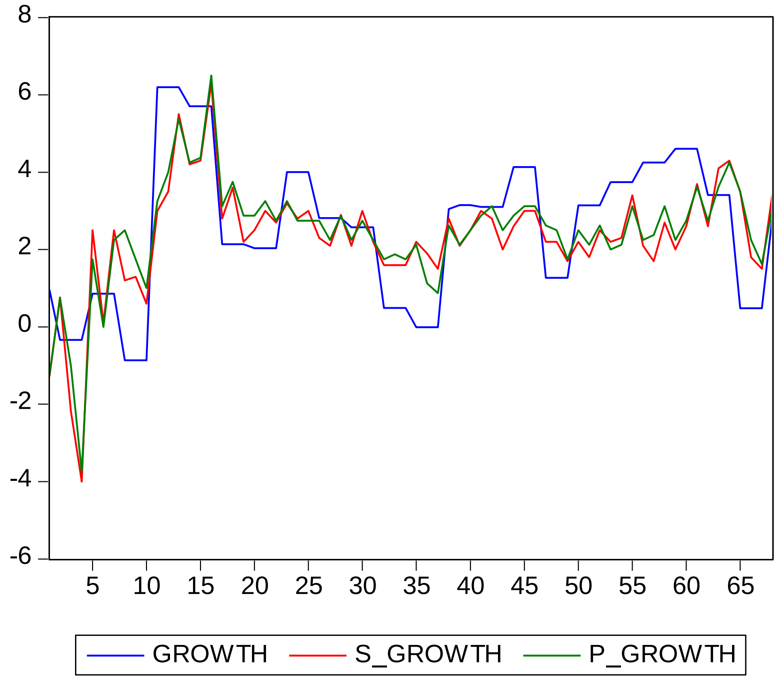

Figure 3

Growth rate, Staff forecasts (S_growth) and FOMC forecasts (P_growth)

The sample is from 1979 to 2001, with 22 February forecasts and 46 July forecasts (1979 is observation 1 and 2001 is observation 68).

References

Batchelor, R. (2007), Bias in Macroeconomic Forecasts, *International Journal of Forecasting*, 23, 189-203.

Chang, C.-L., P.H. Franses and M. McAleer (2011), How Accurate are Government Forecasts of Economic Fundamentals? The Case of Taiwan, *International Journal of Forecasting*, 27, 1066-1075.

Chong, Y.Y. and D.F. Hendry (1986), Econometric Evaluation of Linear Macro-Economic Models, *Review of Economic Studies*, 53, 671-690.

Clark, T.E. and M.W. McCracken (2001), Tests of Equal Forecast Accuracy and Encompassing for Nested Models, *Journal of Econometrics*, 105, 85-110.

Clements, M.P. and D.F. Hendry (2002) (eds.), *A Companion to Economic Forecasting*, Oxford: Blackwell.

Diebold, F.X. and R.S. Mariano (1995), Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

Elliott, G. and A. Timmermann (2008), Economic Forecasting, *Journal of Economic Literature*, 46, 3-56.

Eroglu, C. and K.L. Croxton (2010), Biases in Judgmental Adjustments of Statistical Forecasts: The Role of Individual Differences, *International Journal of Forecasting*, 26, 116-133.

Fiebig, D.G., M. McAleer and R. Bartels (1992), Properties of Ordinary Least Squares Estimators in Regression Models with Non-Spherical Disturbances, *Journal of Econometrics*, 54, 321-334.

Fildes, R, P. Goodwin, M. Lawrence, and K. Nikolopoulos (2009), Effective Forecasting and Judgemental Adjustments: An Empirical Evaluation and Strategies for Improvement in Supply-Chain Planning, *International Journal of Forecasting*, 25, 3-23.

Franses, P.H., H. Kranendonk, and D. Lanser (2011), One Model and Various Experts: Evaluating Dutch Macroeconomic Forecasts, *International Journal of Forecasting*, 27, 482-495.

Franses, P.H. and R. Legerstee (2009), Properties of Expert Adjustments on Model-Based SKU-level Forecasts, *International Journal of Forecasting*, 25, 35-47.

Franses, P.H., M. McAleer and R. Legerstee (2009), Expert Opinion Versus Expertise in Forecasting, *Statistica Neerlandica*, 63, 334-346.

Granger, C.W.J. and P. Newbold (1986), *Forecasting Economic Time Series*, San Diego: Academic Press.

Heij, C., D.J.C. van Dijk and P.J.F. Groenen (2011), Real-time Macroeconomic Forecasting with Leading Indicators: An Empirical Comparison, *International Journal of Forecasting*, 27, 466-481.

McAleer, M. (1992), Efficient Estimation: The Rao-Zyskind Condition, Kruskal's Theorem and Ordinary Least Squares, *Economic Record*, 68, 65-72.

McAleer, M. and C.R. McKenzie (1991), When Are Two Step Estimators Efficient?, *Econometric Reviews*, 10, 235-252.

Romer, C.D. and D.H. Romer (2004), A New Measure of Monetary Shocks: Derivation and Implications, *American Economic Review*, 94, 1055-1084.

Romer, C.D. and D.H. Romer (2008), The FOMC Versus the Staff: Where Can Monetary Policymakers Add Value?, *American Economic Review: Papers & Proceedings*, 98, 230-235.

Smith, J. and M. McAleer (1994), Newey-West Covariance Matrix Estimates for Models with Generated Regressors, *Applied Economics*, 26, 635-640.

Stock, J.H. and M.W. Watson (2003), Forecasting Output aNd Inflation: The Role of Asset Prices, *Journal of Economic Literature*, 41, 788-829.

Timmermann, A. (2006), Forecast Combinations, in G. Elliott, C.W.J. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Amsterdam: North Holland.

West, K.D. (1996), Asymptotic Inference About Predictive Ability, *Econometrica*, 64, 1067-1084.